

#8

**PATENT APPLICATION**

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

10/21/00  
09/697122 PRO  
JC917 U.S.

In re the Application of:

Yukio TABEI

New Application

Filed: October 27, 2000

Attorney Dkt. No.: 32011-167412

For: SPEECH SYNTHESIS DEVICE

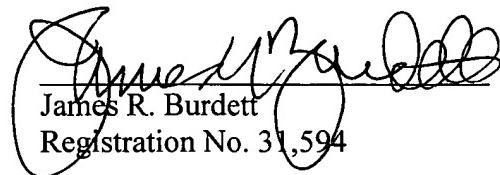
**SUBMISSION OF PRIORITY DOCUMENT**

Assistant Commissioner for Patents  
Washington, D.C. 20231

Sir:

Submitted herewith is a certified copy of Japanese Application No. 2000-075831 upon which a claim to priority was made under 35 U.S.C. §119. The Examiner is respectfully requested to acknowledge receipt of said priority document.

Respectfully submitted,

  
\_\_\_\_\_  
James R. Burdett  
Registration No. 31,594

Venable  
Post Office Box 34385  
Washington, D.C. 20043-9998  
Telephone: (202) 962-4800  
Facsimile: (202) 962-8300

Date: October 27, 2000

JRB:lrh  
#248048

10917 U.S. PRO  
09/697122  
10/27/00

**PATENT OFFICE  
JAPANESE GOVERNMENT**

This is to certify that the annexed is a true copy of the following application as filed with this office.

Date of Application: March 17, 2000

Application Number: 2000-075831

Applicant(s): Oki Electric Industry Co., Ltd.

Dated August 4, 2000

Commissioner,  
Patent Office      Kozo Oikawa

Certificate No. 2000-3060650

日本国特許庁  
PATENT OFFICE  
JAPANESE GOVERNMENT

Jc917 U.S. PRO  
09/697122  
10/12/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されて  
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed  
with this Office.

出願年月日  
Date of Application:

2000年 3月17日

出願番号  
Application Number:

特願2000-075831

出願人  
Applicant(s):

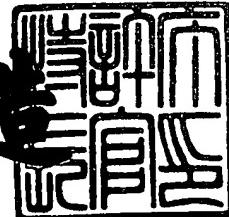
沖電気工業株式会社

CERTIFIED COPY OF  
PRIORITY DOCUMENT

2000年 8月 4日

特許庁長官  
Commissioner,  
Patent Office

及川耕三



出証番号 出証特2000-306065

【書類名】 特許願  
【整理番号】 OG004303  
【提出日】 平成12年 3月17日  
【あて先】 特許庁長官殿  
【国際特許分類】 G10L 3/00  
G10L 5/04

## 【発明者】

【住所又は居所】 東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

【氏名】 田部井 幸雄

## 【特許出願人】

【識別番号】 000000295

【氏名又は名称】 沖電気工業株式会社

## 【代理人】

【識別番号】 100089093

## 【弁理士】

【氏名又は名称】 大西 健治

## 【手数料の表示】

【予納台帳番号】 004994

【納付金額】 21,000円

## 【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9720320

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 規則音声合成装置

【特許請求の範囲】

【請求項1】 予め格納してある音声合成単位を選択して接続し、韻律情報を制御して、任意の音声を合成する規則音声合成装置において、

閉鎖区間を有する音韻の閉鎖区間長を、母音長、子音長とは独立に予測し、制御する音韻継続時間設定手段を備えたことを特徴とする規則音声合成装置。

【請求項2】 請求項1記載の規則音声合成装置において、前記音韻継続時間設定手段は、

入力音韻記号列に対して、音韻の種類を判定する音韻種類判定手段と、母音長予測手段、母音長学習手段、を備えた母音長決定手段と、子音長予測手段、子音長学習手段、を備えた子音長決定手段と、閉鎖長予測手段、閉鎖長学習手段、を備えた閉鎖長決定手段と、を備え、

前記音韻種類判定手段は着目している音韻が母音であるか子音であるかにより、前記母音長予測手段又は子音長予測手段を駆動し、また、子音であると判定された場合には閉鎖長を前に伴うか否かを判定し、閉鎖長を前に伴う場合には閉鎖長予測手段を駆動することを特徴とする規則音声合成装置。

【請求項3】 請求項2記載の規則音声合成装置において、前記閉鎖長決定手段は、更に、閉鎖長分類手段を備え、

該閉鎖長分類手段は、学習データの閉鎖長の度数分布を求め、該度数分布に基づき閉鎖長を第一のグループに分類し、第一のグループに基づき着目音韻を第二のグループに分類する分類動作を行い、

前記閉鎖長学習手段は、前記第二のグループ毎に学習し、学習により得られた音韻継続時間の予測に必要な重み係数を閉鎖長予測手段に送出する学習動作を行い、

前記閉鎖長予測手段は、入力音韻記号列から着目している音韻の音韻名を判定し、該音韻名から前記第二のグループを判定・選択し、前記グループ固有の重み係数を選択し、該重み係数を用いて閉鎖長の予測を行う動作を行い、予測閉鎖長

の値を出力することを特徴とする規則音声合成装置。

**【請求項4】** 請求項2記載の規則音声合成装置において、前記母音長決定手段は、更に、母音長分類手段を備え、

該母音長分類手段は、学習データの母音長の度数分布を求め、該度数分布に基づき母音長を第一のグループに分類し、第一のグループに基づき着目音韻を第二のグループに分類する分類動作を行い、

前記母音長学習手段は、前記第二のグループ毎に学習し、学習により得られた音韻継続時間の予測に必要な重み係数を母音長予測手段に送出する学習動作を行い、

前記母音長予測手段は、入力音韻記号列から着目している音韻の音韻名を判定し、該音韻名から前記第二のグループを判定・選択し、前記グループ固有の重み係数を選択し、該重み係数を用いて母音長の予測を行う動作を行い、予測母音長の値を出力することを特徴とする規則音声合成装置。

**【請求項5】** 請求項2記載の規則音声合成装置において、前記子音長分類手段は、更に、子音長分類手段を備え、

該子音長分類手段は、学習データの子音長の度数分布を求、該度数分布に基づき子音長を第一のグループに分類し、第一のグループに基づき着目音韻を第二のグループに分類する分類動作を行い、

前記子音長学習手段は、前記第二のグループ毎に学習し、学習により得られた音韻継続時間の予測に必要な重み係数を子音長予測手段に送出する学習動作を行い、

前記子音長予測手段は、入力音韻記号列から着目している音韻の音韻名を判定し、該音韻名から前記第2のグループを判定・選択し、前記グループ固有の重み係数を選択し、該重み係数を用いて子音長の予測を行う動作を行い、予測子音長の値を出力することを特徴とする規則音声合成装置。

**【請求項6】** 請求項3記載の規則音声合成装置において、

前記閉鎖長学習手段は、着目音韻、該着目音韻の前後2音韻からなる音韻環境、音韻位置、品詞等からなる要因を抽出して数量化する第一の要因抽出手段と、直前の音韻が無声化しているか否かを学習データに基づき判定する第一の前方無

声化判定手段と、前記分類された第二のグループ毎に各要因に対する重み係数を作成するモデル学習手段と、から構成され、

前記閉鎖長予測手段は、着目音韻、該着目音韻の前後2音韻からなる音韻環境、音韻位置、品詞等からなる要因を抽出して数量化する第二の要因抽出手段と、所定の無声化規則に基づき着目する音韻の無声化の可否を判定する第2の前方無声化判定手段と、当該音韻から前記第二のグループを判定し、該グループ毎に前記モデル学習手段から出力される重み係数を参照して閉鎖長を予測することを特徴とする規則音声合成装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、音声合成、特に任意の語彙を音声合成する規則音声合成装置に関するものである。

【0002】

【従来の技術】

従来、テキスト文章を音声にして出力するテキスト音声変換は、テキスト解析部と規則音声合成部（パラメータ生成部と音声合成部）から構成される。

【0003】

テキスト解析部では、漢字かな混じり文を入力して、単語辞書を参照して形態素解析し、（必要なら構文解析、意味解析等も行って）、読み、アクセント、イントネーションを決定し、韻律記号付き発音記号（中間言語）を出力する。

【0004】

パラメータ生成部では、ピッチ周波数パターンや音韻継続時間長、ポーズ、振幅等の設定を行う。

【0005】

音声合成部では、目的とする音韻系列（中間言語）から音声合成単位を、あらかじめ蓄積されている音声データから選択し、パラメータ生成部で決定したパラメータに従って、結合／変形して音声の合成処理を行う。音声合成単位は、音素、音節（C V）、V C V, C V C (C:子音、V:母音) などが試みられてきた

。音素は、最も数が少なく表現できるが、調音結合に対する規則化が不可欠であり、その規則化が困難なため、音質は悪く、現在ではほとんど用いられていない。C V, V C V, C V C の単位は、単位の中に調音結合を含み、V C V は母音で子音をはさむため、子音の明瞭度が高く、C V C は振幅の小さい子音で接続するため接続歪みは小さい。また最近は、これらの音韻連鎖を拡張した単位も一部用いられている。

## 【0006】

音声合成単位データの表現方法としては、原音声波形をそのまま利用して、品質劣化の少ない高品質の合成音を得る手法が用いられるようになって来ている。

## 【0007】

以上説明した構成のテキスト音声変換によって、より自然性の高い合成音声を出力するには、音声合成単位の種類、素片品質、合成方式と共に、前記パラメータ生成部でのパラメータ（ピッチ周波数パターン、音韻継続時間長、ポーズ、振幅）をいかに自然音声に近くなるよう適切に制御するかがきわめて重要となる。

## 【0008】

それらのパラメータの中で、特に、音韻継続時間長を制御する方法としては、従来、文献1（特開昭63-46498）、文献2（特開平4-134499）に記載される方法がある。

## 【0009】

上記文献1、2に記載された技術は、統計数理的モデル（数量化1類モデル）を用いて、多量のデータを解析して、制御規則を求める方法である。数量化1類モデルは、公知のように、多変量解析の1つであり、質的な要因に基づいて目的となる外的基準（音韻継続時間長）を算出するもので、以下の式（1）～（3）により定式化される。

## 【0010】

すなわち、 $i$  番目のデータの要因アイテムを  $j$ 、その属するカテゴリを  $k$ 、そのカテゴリ数量（カテゴリに付与する係数）を  $x(j\ k)$  とするとき、予測値  $y(i)$  は式（1）となる。

$$y(i) = \sum_{j,k} x(j,k) \delta(j,k) \quad \dots (1)$$

ここで、

$$\begin{aligned} \delta(j,k) &= 1 \text{ (データ } i \text{ が } j \text{ アイテムの } k \text{ カテゴリに反応した時)} \\ &= 0 \text{ (それ以外)} \end{aligned} \quad \dots (2)$$

#### 【0011】

$x(j,k)$  は、最小2乗法で求められる。すなわち、予測値  $y(i)$  と実測値  $Y(i)$  の2乗誤差が最小になるようにして求められる。

$$\sum_i \{ y(i) - Y(i) \}^2 \quad \rightarrow \text{最小} \quad \dots (3)$$

#### 【0012】

式(3)を  $x(j,k)$  で偏微分して方程式を解く必要があり、コンピュータによる実際の計算としては、連立方程式を解く数値解析問題に帰着できる。

#### 【0013】

##### 【発明が解決しようとする課題】

上述の、従来の音韻継続時間長制御方法では、数量化 I 類でカテゴリ化がうまく行なえないことがあり、十分な予測精度を達成できないことがあった。また、これらの従来の方法では閉鎖区間を有する音韻（無声破裂音など）に対しては、その閉鎖長の決定方法については何ら記載が無く、知覚上大切な閉鎖区間長を適切に制御する方法が存在していなかった。

#### 【0014】

本発明は、音韻継続時間長の予測精度を上げ、推定誤差を小さくし、制御性能を向上させるもので、特に、閉鎖区間を有する音韻（無声破裂音など）に対して、その閉鎖長の適切な制御方法を提供することを主眼とし、その結果、品質を向上させた規則音声合成装置を提供することを目的とする。

#### 【0015】

##### 【課題を解決するための手段】

そのために、本発明の規則音声合成装置においては、予め格納してある音声合成単位を選択して接続し、韻律情報を制御して、任意の音声を合成する規則音声

合成装置において、閉鎖区間を有する音韻の閉鎖区間長を、母音長、子音長とは独立に予測し、制御する音韻継続時間設定手段を備えたことを特徴とする。

#### 【0016】

##### 【発明の実施の形態】

以下、本発明の実施の形態について、図面を参照しながら詳細に説明する。

##### ＜音声合成装置の基本的な構成＞

図1は、本発明の実施形態に於ける音声合成装置（テキスト音声変換装置）の構成図を示したもので、テキスト解析部101では、漢字かな混じり文を入力して、単語辞書102を参照して形態素解析し、読み、アクセント、イントネーションを決定し、韻律記号付き発音記号（中間言語）を出力する。

#### 【0017】

パラメータ生成部103では、中間言語自身から使用すべき、素片辞書105内の素片アドレスを選択し、また、ピッチ周波数パターンや音韻継続時間長、振幅等の設定を行う。

#### 【0018】

素片辞書105は、音声信号を入力した後、あらかじめ素片作成部106にて作成される。

#### 【0019】

素片作成部106では、音声合成する前に、あらかじめ音声データから合成音の基となる素片を作成しておく。

#### 【0020】

音声合成部104は、従来の種々の方法が適用でき、例えば、波形重畠法を用いることができる。なお、韻律記号付き発音記号（中間言語）を入力として音声合成を行なうのが、規則音声合成である。

#### 【0021】

パラメータ生成部103で決定した音韻の継続時間長は、日本語の等モーラ規則に基づき、主に母音部の伸縮によって音韻継続時間長を実現する。すなわち、決定した音韻継続時間が素片より長い場合は、最後尾の素片を繰り返し使用し（伸長）、短い場合は、途中で打ち切る（圧縮）処理を行なう。

## 【0022】

尚、図1において、テキスト解析部101、単語辞書102、音声合成部104、素片辞書105、素片作成部106、は従来の技術を用いて構成できる。

## 【0023】

<パラメータ生成部における音韻継続時間長設定方法の第1の実施形態>

パラメータ生成部103における音韻継続時間長設定方法の第1の実施形態について、図2を参照して詳細に説明する。

## 【0024】

図において、音韻記号列が入力され、音韻種類判定部201において、着目している音韻の種類が、母音であるか、子音であるか、子音と判定された場合には、閉鎖長を時間的に前方に伴う子音（／p, t, k／など図6参照）かどうかを判定する。その結果、母音と判定された場合には、母音長予測部202を駆動し、子音と判定された場合には、子音長予測部205を駆動し、更に、／p, t, k／等の当該音韻が音韻の種類によって閉鎖長を前方に伴うと判定された場合には、閉鎖長予測部208を駆動し、それぞれ時間長を予測する。その後、それぞれ母音長設定部203、子音長設定部206、閉鎖長設定部209によって予測された時間長を設定する。子音長の設定の時間的な順は予測閉鎖長、予測子音長で行なう。なお、子音の中で閉鎖長を時間的に前方に伴う子音の種類としては、実際の音声データを解析した結果、図6に示す音韻のみであり、鼻音などは伴わなかった。

## 【0025】

なお、時間長の予測には、例えば、数量化I類による方法を用い、予め学習データ211によって、母音長学習部204、子音長学習部207、閉鎖長学習部210の各モデルを学習し（前述の式（3）のような規準で連立方程式を解くことに相当）、予測に必要な重み係数を決定しておく。この重み係数の決定は学習データを用いて前述の式（1）のx(j k)を連立方程式から決定することである。

## 【0026】

以上説明したように、本実施形態の音韻継続時間長設定方法によれば、閉鎖長

を時間的に前方に伴う音韻に対して、適切な音韻継続時間長を制御することが可能となり、規則音声合成装置において自然性の高い合成音を得ることが可能となる。

#### 【0027】

尚、本実施形態では、学習・予測に数量化I類を用いる構成としたが、これに限定されるものではなく、他の統計的手法を用いても良い。

#### 【0028】

<パラメータ生成部における音韻継続時間長設定方法の第2の実施形態>

パラメータ生成部103における音韻継続時間長設定方法の第2の実施形態について、図3を参照して詳細に説明する。

#### 【0029】

図において、閉鎖長分類部301を設けた点、および閉鎖長学習部302と閉鎖長予測部303の動作が第1の実施形態と異なり、第1の実施形態と同様の動作の個所は、図2と同一の番号を付与してある。以下、その動作について説明する。

#### 【0030】

まず、音韻記号列が入力され、音韻種類判定部201において、着目している音韻の種類が、母音であるか、子音であるか、子音と判定された場合には、閉鎖長を時間的に前方に伴うかどうかを判定する。その結果、母音と判定された場合には、母音長予測部202を駆動し、子音と判定された場合には、子音長予測部205を駆動し、更に、閉鎖長を前方に伴うと判定された場合には、閉鎖長予測部303を駆動し、それぞれ時間長を予測する。その後、それぞれ母音長設定部203、子音長設定部206、閉鎖長設定部209によって予測された時間長を設定する。子音長の設定の時間的な順は予測閉鎖長、予測子音長の順で行なう。

#### 【0031】

尚、時間長の予測には数量化I類による方法を用い、数量化I類モデルを用いて閉鎖長を学習／予測する方法が第1の実施形態とは異なる。すなわち、図3において、予め学習データ211を閉鎖長分類部301によって分類し、閉鎖長学習部302の各モデルを学習し、予測に必要な重み係数を決定しておく。

## 【0032】

数量化I類は、カテゴリ数の数だけの線形重み和でモデル化を行なうため、学習データの信頼性によって推定精度が決まってしまう。また、要因としては、着目している音韻、前後2つの音韻環境、音韻の位置などを用いるが、一般にこれらの要因は質的データであり大小順には並んでいない。このため本質的に要因をグループ分けすることはできない。

## 【0033】

第2の実施形態では、この点を改善するためのものであり、本実施形態の特徴である閉鎖長分類部301、閉鎖長学習部302、閉鎖長予測部303の動作を図7を用いて説明する。

## 【0034】

図7において、閉鎖長分類部301では、ステップ701で学習データの外的基準（閉鎖長）の度数分布を求める。ステップ702で度数分布をもとにして、いくつかのグループに分け、ステップ703で着目している音韻との対応をとりこの音韻もグループ分けを行う。

## 【0035】

閉鎖長学習部302では、ステップ704で前述のグループ毎に学習を行ない、重み係数を学習し、ステップ705で閉鎖長予測部303へ送出する。

## 【0036】

予測の際には、閉鎖長予測部303では、ステップ710で入力音韻記号列から当該音韻名を判定し、ステップ711で当該音韻名から前記グループを判定・選択して、ステップ712で、前記グループ固有の重み係数を選択し、ステップ713で前記重み係数を用いて数量化I類で閉鎖長の予測を行なう。

## 【0037】

以上説明したように、本実施形態の音韻時間長設定方法によれば、閉鎖長を前述のようにグループに分類することにより、実際にあらわれる閉鎖長の分布を的確にとらえることが可能となり、学習が従来の方法より精度よく行なえ、予測においては、予測値の分散が小さく抑えられ、予測精度が向上する効果がある。

## 【0038】

## &lt;パラメータ生成部における音韻継続時間長設定方法の第3の実施形態&gt;

パラメータ生成部103における音韻継続時間長設定方法の第3の実施形態について、図4を参照して詳細に説明する。

## 【0039】

図において、母音長分類部401、子音長分類部404を設けた点、および母音長学習部402、母音長予測部403、子音長学習部405、子音長予測部406の動作が第2の実施形態と異なり、第2の実施形態と同じ動作の個所は、図3と同一の番号を付与してある。以下、その動作について説明する。

## 【0040】

まず、音韻記号列を入力し、音韻種類判定部201において、現在着目している音韻の種類が、母音であるか、子音であるか、子音と判定された場合には、閉鎖長を時間的に前方に伴うかどうかを判定する。その結果、母音と判定された場合には、母音長予測部403を駆動し、子音と判定された場合には、子音長予測部406を駆動し、閉鎖長を前方に伴うと判定された場合には、閉鎖長予測部303を駆動し、それぞれ時間長を予測する。その後、それぞれ母音長設定部203、子音長設定部206、閉鎖長設定部209によって予測された時間長を設定する。子音長の設定の時間的な順は予測閉鎖長、予測子音長の順で行なう。

## 【0041】

図4で、予め学習データ211のうち母音長学習データを母音長分類部401によって、子音長学習データを子音長分類部404によって分類する。また、閉鎖長に関しては、閉鎖長学習データを閉鎖長分類部301によって分類し、閉鎖長学習部302、閉鎖長予測部303を動作させるのは第2の実施形態と同一であるため、説明を省略する。

## 【0042】

数量化I類の要因は質的データであり大小順には並んでいない。このため本質的に要因をグループ分けすることはできない。第3の実施形態では、第2の実施形態と同様に、この点を改善するためのものであるが、特に母音長、子音長の予測精度を改善するものである。

## 【0043】

第3の実施形態の特徴である母音長分類部401、母音長学習部402、母音長予測部403の動作を図8に、また、子音長分類部404、子音長学習部405、子音長予測部406の動作を図9に示す。

#### 【0044】

母音長に関しては図8において、ステップ801で学習データの外的基準（母音長）の度数分布を求める。ステップ802で度数分布をもとにして、いくつかのグループに分け、ステップ803で当該音韻との対応をとり当該音韻をもグループ分けを行う。母音長学習部402では、ステップ804で前記グループ毎に学習を行ない、重み係数を学習し、ステップ805で母音長予測部403へ送出する。

#### 【0045】

母音長予測部403における予測の際には、ステップ810で入力音韻記号列から当該音韻名を判定し、ステップ811で当該音韻名から前記グループを判定・選択して、ステップ812で、前記グループ固有の重み係数を選択し、ステップ813で前記重み係数を用いて数量化I類で母音長の予測を行なう。

#### 【0046】

同様に、子音に関しては図9において、ステップ901で学習データの外的基準（子音長）の度数分布を求める。ステップ902で度数分布をもとにして、いくつかのグループに分け、ステップ903で当該音韻との対応をとり当該音韻もグループ分けを行う。子音長学習部405では、ステップ904で前記グループ毎に学習を行ない、重み係数を学習し、ステップ905で子音長予測部406へ送出する。

#### 【0047】

子音長予測部406における予測の際には、ステップ910で入力音韻記号列から当該音韻名を判定し、ステップ911で当該音韻名から前記グループを判定・選択して、ステップ912で、前記グループ固有の重み係数を選択し、ステップ913で前記重み係数を用いて数量化I類で子音長の予測を行なう。

#### 【0048】

以上説明したように、本実施形態によれば、母音長、子音長は、単純な分布で

はなく、一般には多峰性分布をしていて、上記のようにグループに分類することにより、従来に比べ、学習データを的確に捕らえた学習が可能になり、予測においては、予測値の平均値が前記グループの平均値となり、予測値の分散が小さく抑えられ、予測精度が向上する効果がある。

## 【0049】

## &lt;パラメータ生成部における音韻継続時間長設定方法の第4の実施形態&gt;

パラメータ生成部103における音韻継続時間長設定方法の第4の実施形態について、図5を参照して詳細に説明する。

## 【0050】

図において図2、図3と同一機能のブロックは同一番号を記してある。同図において閉鎖長予測部208は、要因抽出部501、前方無声化判定手段502、予測モデル部503から構成され、閉鎖長学習部210は、要因抽出部505、前方無声化判定手段506、学習モデル部504から構成される。それらの動作について、以下説明する。

## 【0051】

まず、学習データ211のうちの閉鎖長学習データ510を、第2の実施形態と同様に閉鎖長分類部303によってグループの分類を行なう。その後、要因抽出部505により、当該音韻、前後2つの音韻環境、音韻位置（呼気段落内、文内）、モーラ数（呼気段落、文）、品詞などの要因を抽出し数量化し、学習モデル部504へ供給する。同時に、前方無声化判定手段506によって直前の音韻が無声化しているかどうかを学習データに基づいて判定し、無声化している場合1、していない場合2の数値データを生成し学習モデル部504へ供給する。学習モデル部504は、数量化I類モデルによって構成され、前記グループ毎に学習結果として、各要因に対する重み係数テーブル520を作成し、予測モデル部503へ重み係数テーブル520を送る。

## 【0052】

予測時には、入力音韻記号列から要因抽出部501において、閉鎖長学習部210における要因抽出部505と同一の要因を抽出し数量化する。それと同時に、前方無声化判定手段502において、後述する無声化規則を適用して音韻の無

声化を決定し、当該音韻の直前の音韻が無声化すべきと決定されたら1、すべきでないと決定されたら2の数値データを生成する。予測モデル部503では、当該音韻から前記グループを判定し、前記グループ毎に重み係数テーブル520を参照し、数量化I類モデルによって閉鎖長を予測する。

#### 【0053】

ここで無声化規則とは、

(1) 無声子音にはさまれた／i／、／u／は無声化する。

ただし、

(2) アクセントがあれば無声化しない。

(3) 連続して無声化しない。

(4) 同じ種類の無声摩擦音にはさまれた母音は無声化しない。

などの規則であり、入力音韻記号列を解析して適用する。

#### 【0054】

以上説明したように、本実施形態によれば、前音韻が無声化しているかどうかで、閉鎖長を制御するため、例えば、「お近く」／o ch i k a k u／の場合、／ch i／の／i／が無声化するので、後続する／ka／の／k／の前に伴う閉鎖区間長を適切な値に制御することが可能になる。

#### 【0055】

なお、第4の実施形態では、前方無声化判定手段502において、後述する無声化規則を適用して音韻の無声化を決定する構成としているが、別の実施形態として、無声化規則を適用するのを、あらかじめ別に行っておき、閉鎖長予測部208では、すでに決定した無声化の情報をもらうような構成にしても何ら差し支えない。

#### 【0056】

##### 【発明の効果】

以上、詳細に説明したように、本発明によれば、予め格納してある音声合成単位を選択して接続し、韻律情報を制御して、任意の音声を合成する規則音声合成装置において、閉鎖区間を有する音韻の閉鎖区間長を、母音長、子音長とは独立に予測し、制御する音韻継続時間設定手段を備えた構成としたので、閉鎖長を時

間的に前方に伴う音韻に対して、適切な音韻継続時間長を制御することが可能となり、規則音声合成装置において自然性の高い合成音を得ることが可能となる。

【図面の簡単な説明】

【図1】

音声合成装置（テキスト音声変換装置）のブロック図である。

【図2】

第1の実施形態における音韻継続時間長設定部の構成図である。

【図3】

第2の実施形態における音韻継続時間長設定部の構成図である。

【図4】

第3の実施形態における音韻継続時間長設定部の構成図である。

【図5】

第4の実施形態における音韻継続時間長設定部の構成図である。

【図6】

閉鎖長を前方に伴う子音の種類を示す図である。

【図7】

第2の実施形態における閉鎖長分類部301、閉鎖長学習部302、閉鎖長予測部303の動作説明図である。

【図8】

第3の実施形態における母音長分類部401、母音長学習部402、母音長予測部403の動作説明図である。

【図9】

第3の実施形態における子音長分類部404、子音長学習部405、子音長予測部406の動作説明図である。

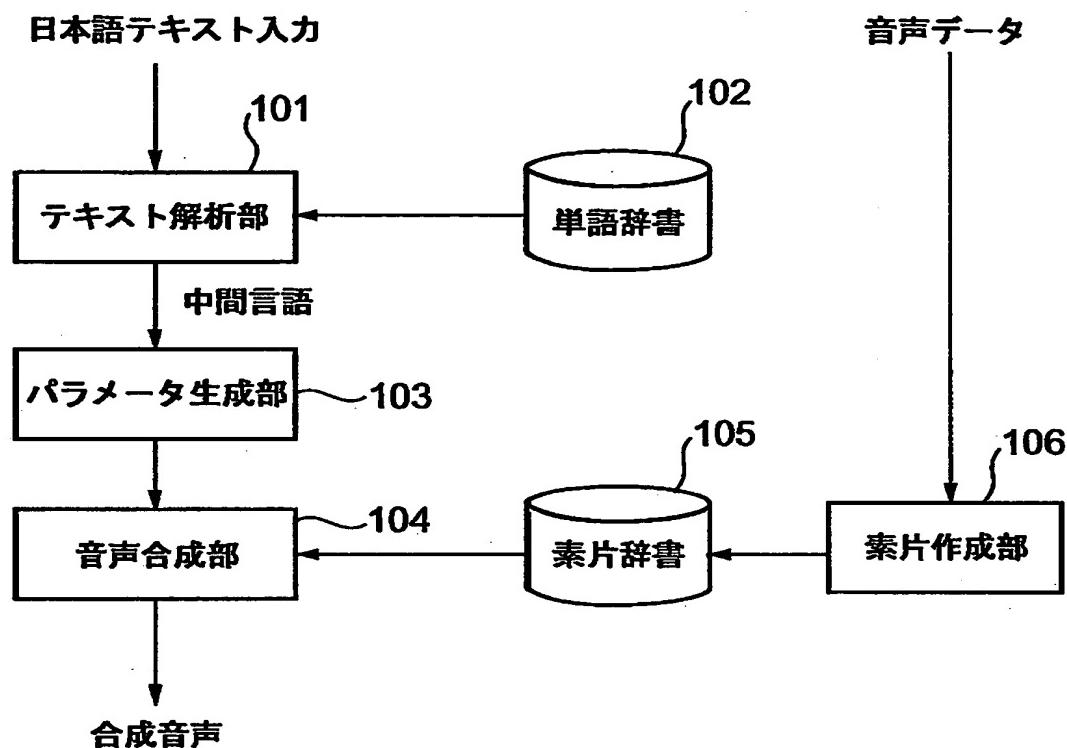
【符号の説明】

101	テキスト解析部
102	単語辞書
103	パラメータ生成部
104	音声合成部

105	素片辞書
106	素片作成部
201	音韻種類判定部
202、403	母音長予測部
203	母音長設定部
204、402	母音長学習部
401	母音長分類部
205、406	子音長予測部
206	子音長設定部
207、405	子音長学習部
404	子音長分類部
208、303	閉鎖長予測部
209	閉鎖長設定部
210、302	閉鎖長学習部
301、303	閉鎖長分類部
211	学習データ
510	閉鎖長学習データ

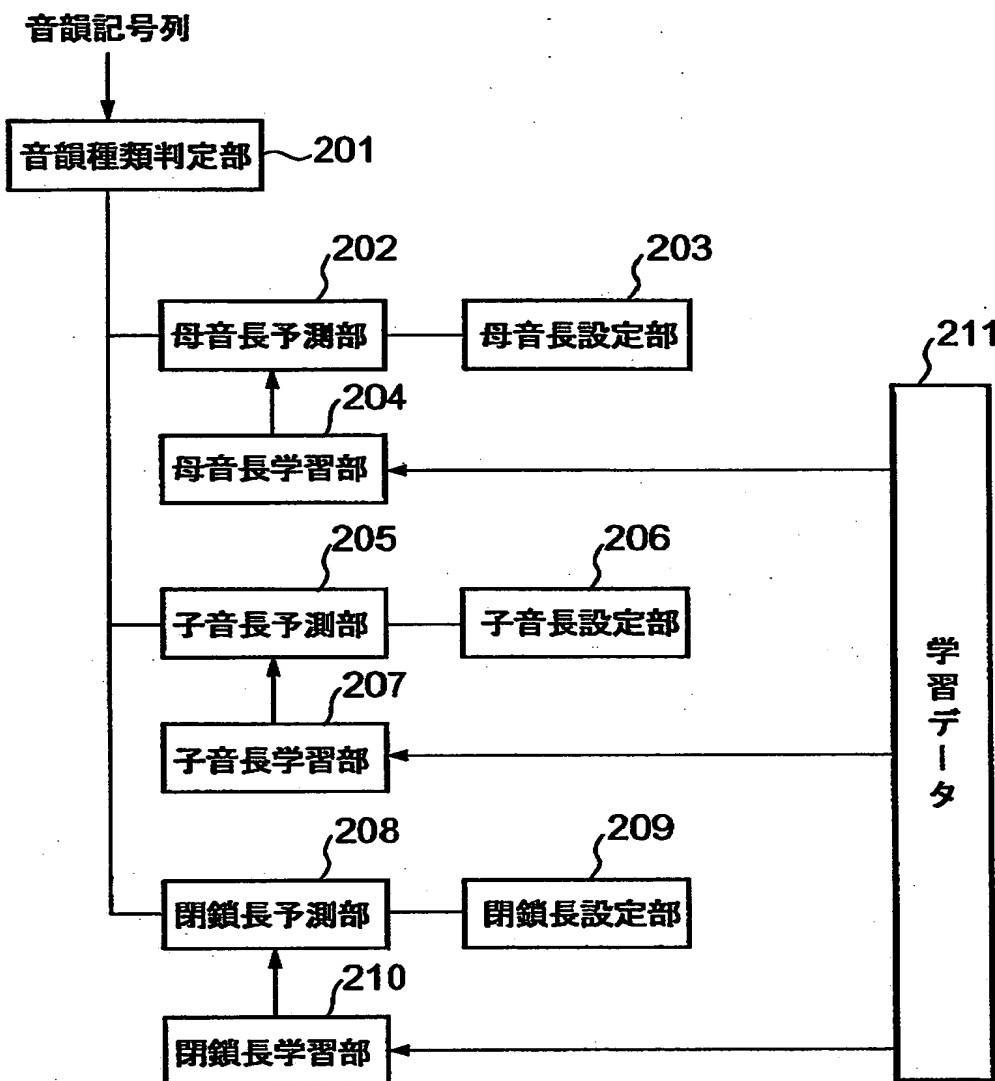
【書類名】 図面

【図1】



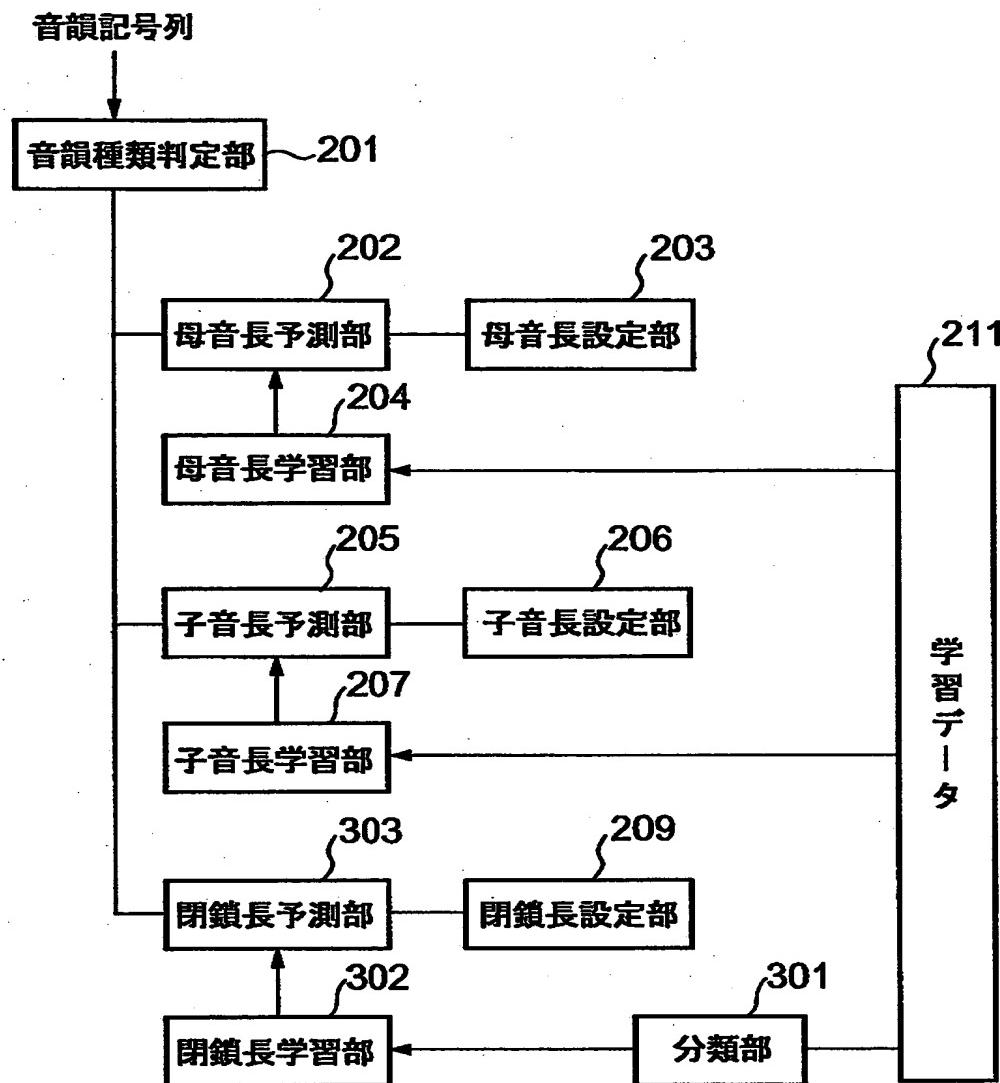
音声合成装置（テキスト音声変換装置）のブロック図

【図2】



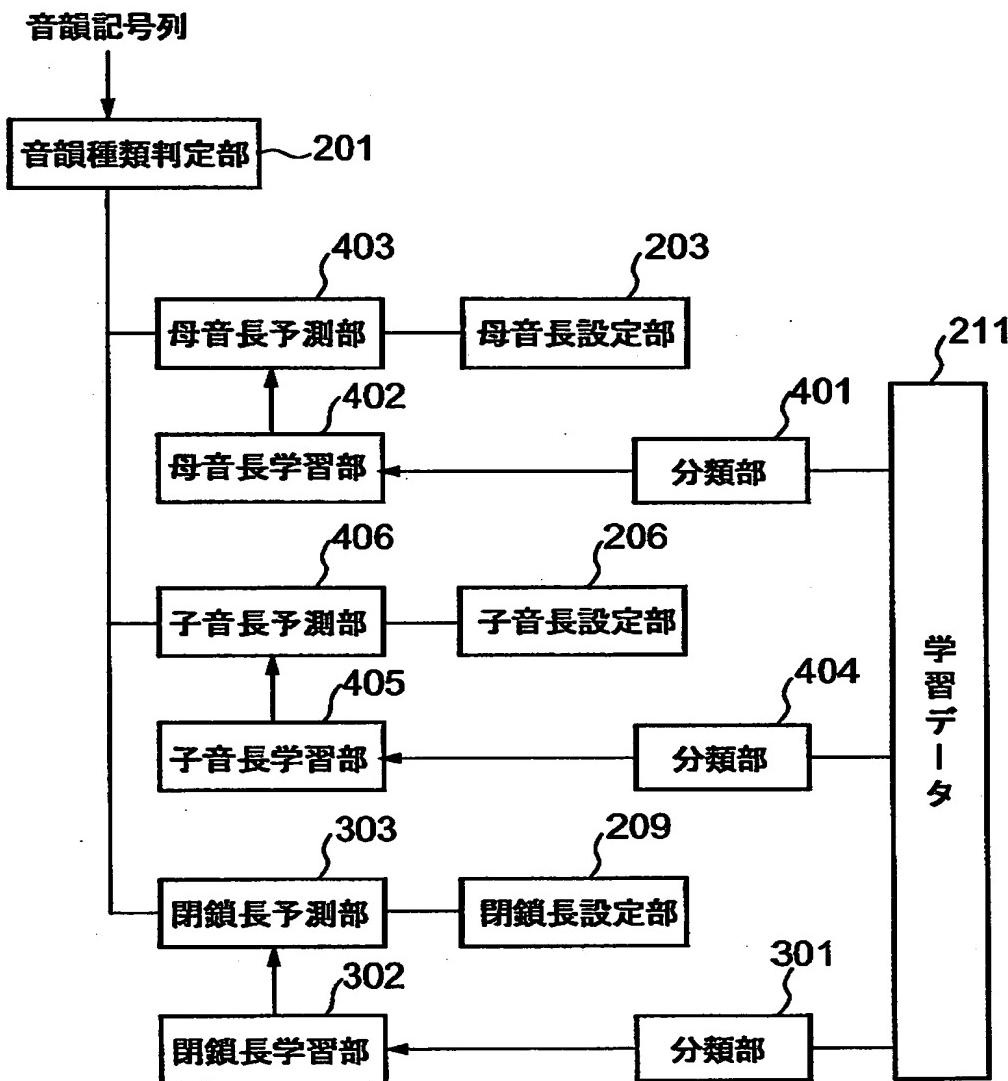
第1の実施形態における音韻継続時間長設定部の構成図

【図3】



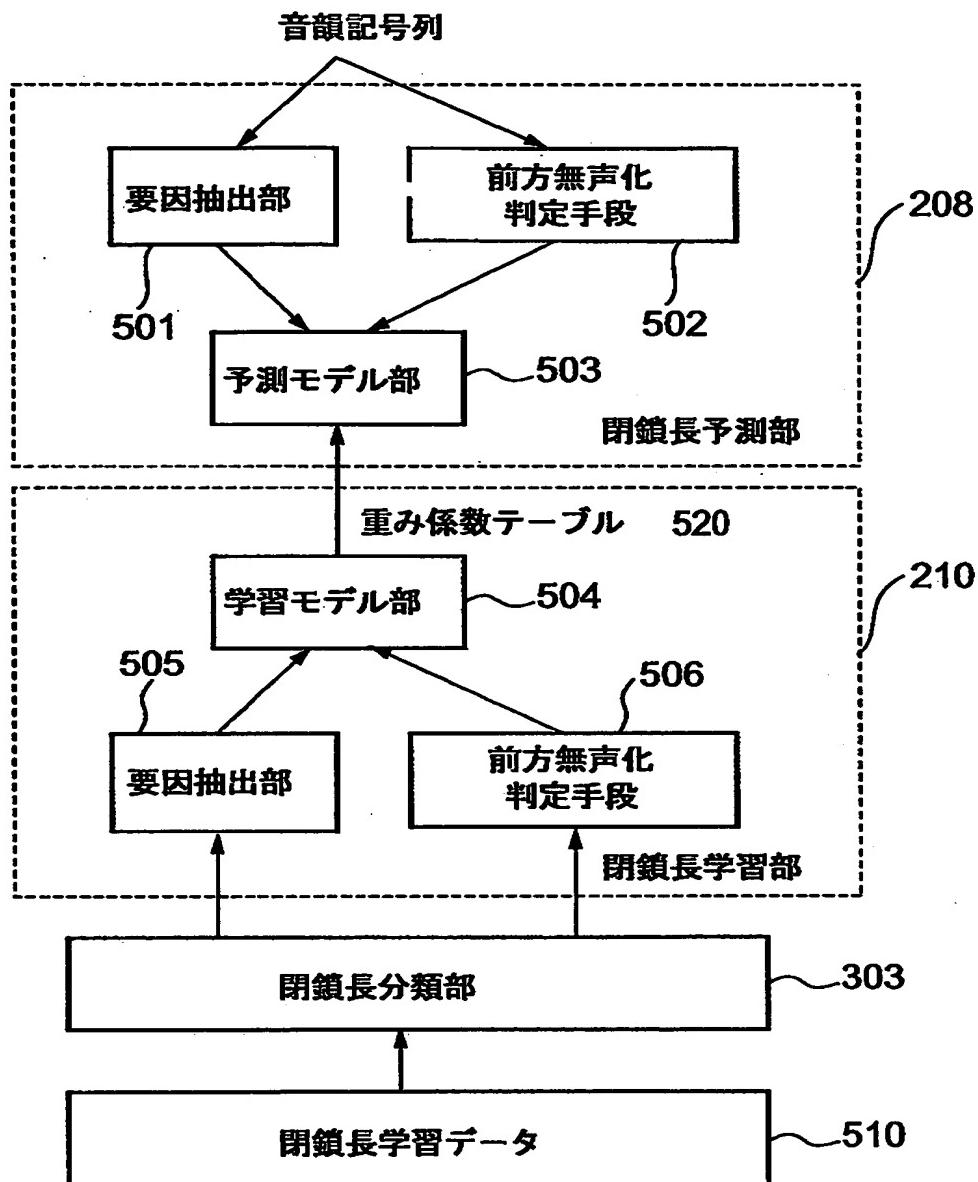
第2の実施形態における音韻継続時間長設定部の構成図

【図4】



第3の実施形態における音韻継続時間長設定部の構成図

【図5】



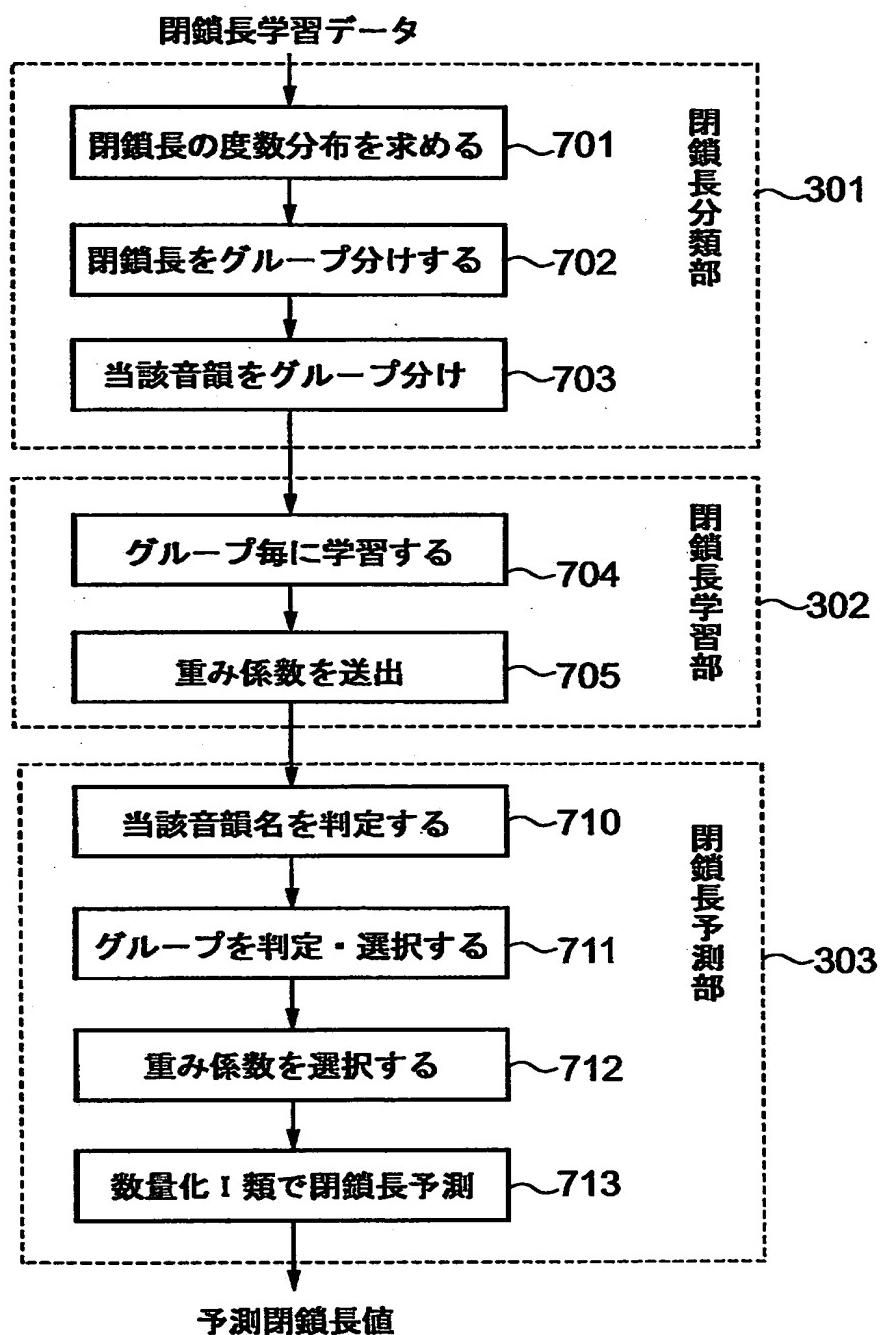
第4の実施形態における閉鎖長学習部210、  
閉鎖長予測部208の構成図

【図6】

番号	種類名	音韻表記
1	無声破裂音	/p, t, k/
2	拗音の一部	/py, ky, cy, /
3	その他	/ts, ch/

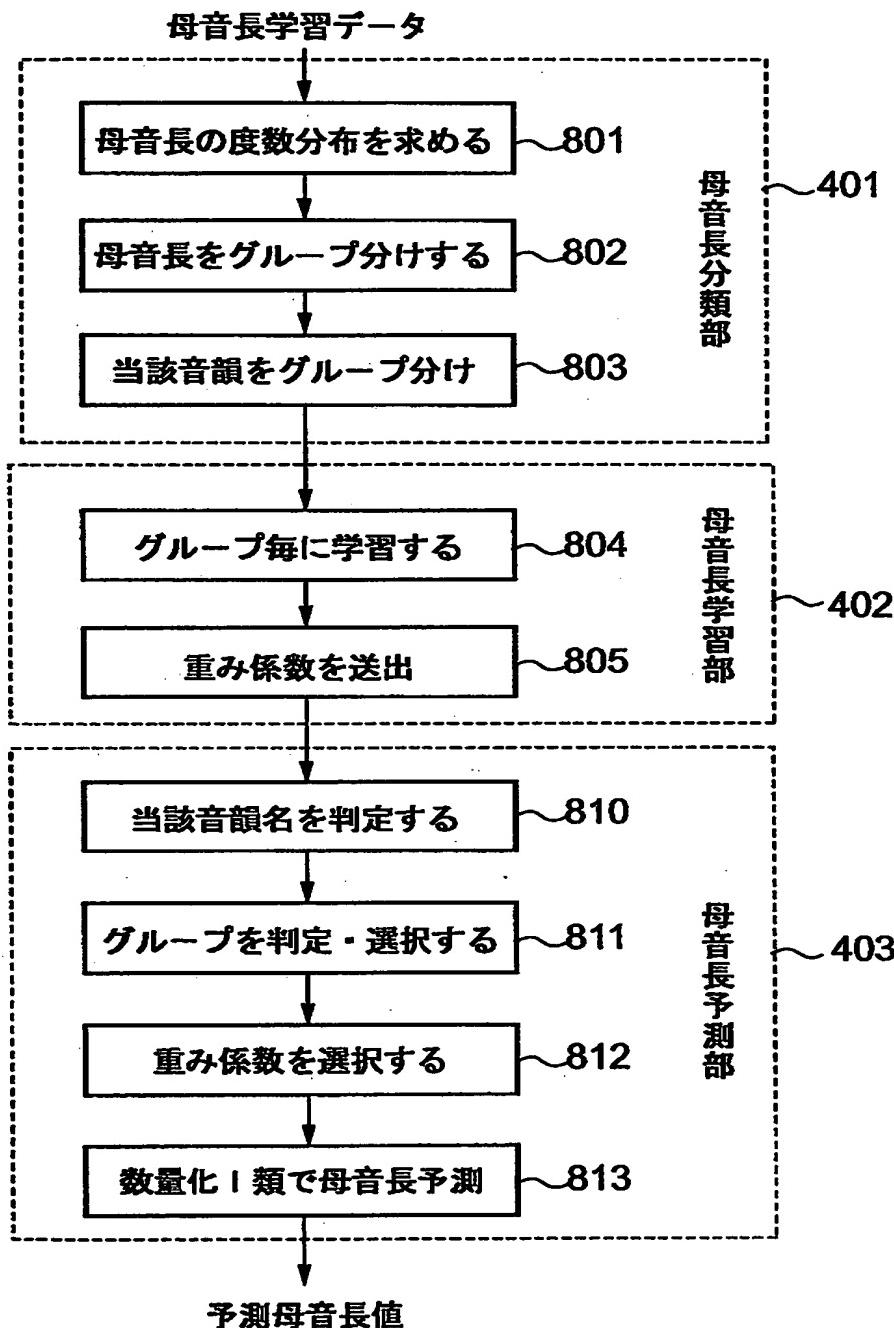
閉鎖長を前方に伴う子音の種類を示す図

【図7】



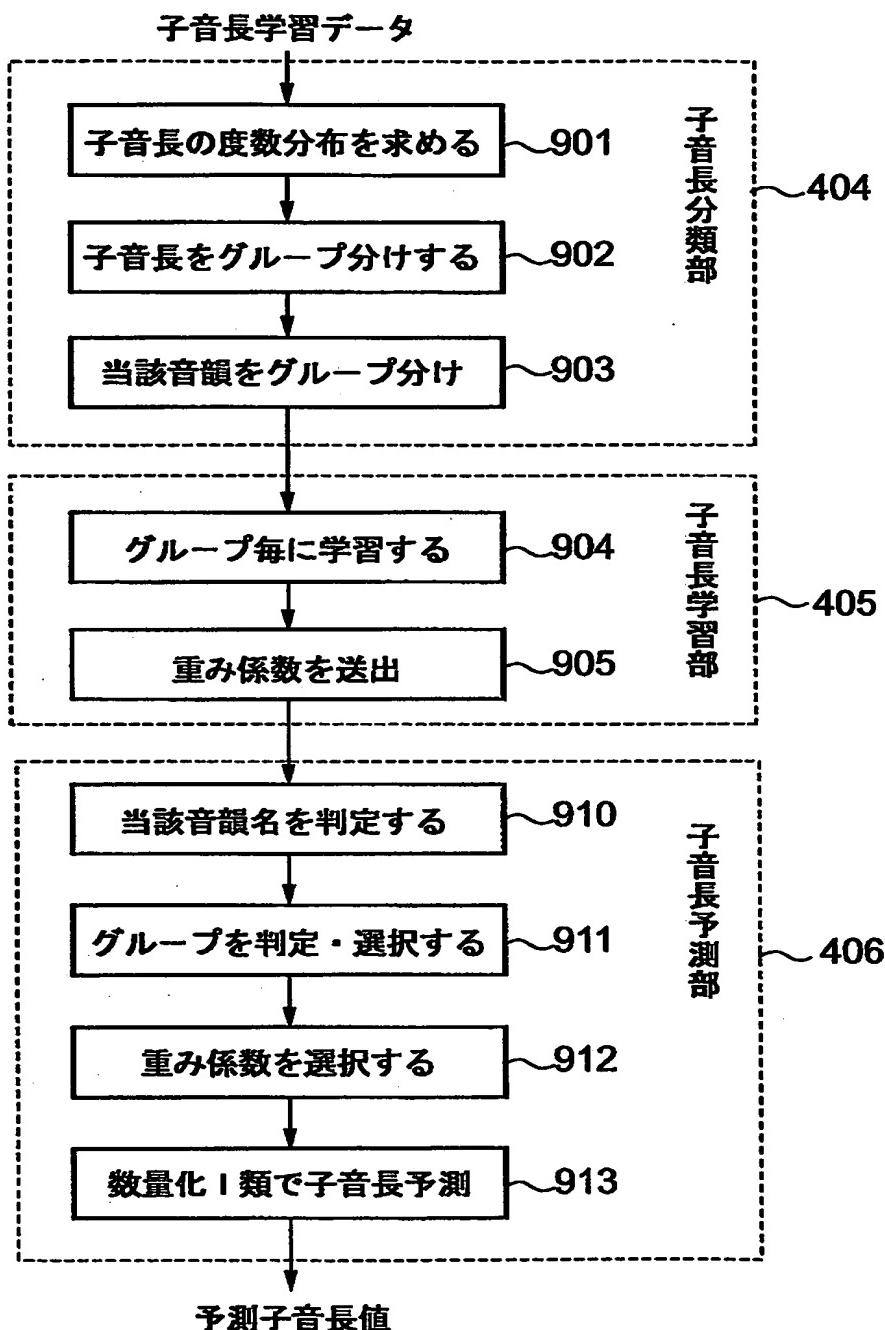
第2の実施形態における閉鎖長分類部301、閉鎖長学習部302、  
閉鎖長予測部303の動作説明図

【図8】



第3の実施形態における母音長分類部401、母音長学習部402、母音長予測部403の動作説明図

【図9】



第3の実施形態における子音長分類部404、子音長学習部405、子音長予測部406の動作説明図

【書類名】 要約書

【要約】

【課題】 閉鎖区間を有する音韻（無声破裂音など）に対して、その閉鎖長の適切な制御方法を提供することを主眼とし、その結果、品質を向上させた規則音声合成装置を提供する。

【解決手段】 音韻種類判定部201において、着目している音韻の種類が、母音であるか、子音であるか、子音と判定された場合には、閉鎖長を時間的に前方に伴う子音かどうかを判定する。その結果、母音と判定された場合には、母音長予測部202を駆動し、子音と判定された場合には、子音長予測部205を駆動し、更に、閉鎖長を前方に伴うと判定された場合には、閉鎖長予測部208を駆動し、それぞれ時間長を予測する。その後、それぞれ母音長設定部203、子音長設定部206、閉鎖長設定部209によって予測された時間長を設定する。

【選択図】 図2

出願人履歴情報

識別番号 [000000295]

1. 変更年月日 1990年 8月22日

[変更理由] 新規登録

住 所 東京都港区虎ノ門1丁目7番12号

氏 名 沖電気工業株式会社